Intrinsic Dirac Behavior of Scalar Curvature in a Complex Weyl-Cartan Geometry

J.E. Rankin*
Rankin Consulting, Concord, CA
(Dated: January 18, 2011)

The "spin-up" and "spin-down" projections of the second order, chiral form of Dirac Theory are shown to fit a superposition of forms predicted in an earlier classical, complex scalar gauge theory[1]. In some sense, it appears to be possible to view the two component Dirac spinor as a single component, quaternionic, spacetime scalar. "Spin space" transformations become transformations of the internal quaternion basis. Essentially, quaternionic Dirac Theory projects into the complex plane neatly, where spin becomes related to the self-dual antisymmetric part of the metric. The correct Dirac eigenvalues and well-behaved eigenfunctions project intact into a pair of complex solutions for the scalar curvature in the earlier theory's Weyl-Cartan type geometry. Some estimates are made for predicted, interesting atomic and subatomic scale phenomena. A generalization of the complex geometric structure to allow quaternionic gauges and curvatures is sketched in an appendix, and appears to be a fairly well defined possibility, but Dirac Theory seems to fit more naturally into the complex plane than into the full, generalized structure.

PACS numbers: 04.20.Cv,04.50.+h,04.20.Fy,03.65.Pm

I. INTRODUCTION

This work presents the conclusion of an extended effort to identify spin 1/2, quantum mechanical wavefunctions with the scalar curvature of a Weyl-like Cartan geometry with a self-dual antisymmetric part to the metric[1–3]. The earlier papers demonstrated that the natural geometric identifications made therein imply as a simplest case that in the limit that the spacetime is Lorentzian ($\hat{g}_{\mu\nu} \approx \hat{\eta}_{\mu\nu}$), there exists a geometric wavefunction ψ which obeys

$$\hat{\eta}^{\mu\nu}\left[\psi_{,\mu,\nu} + \imath q A_{\mu,\nu}\psi + 2\imath q A_\mu \psi_{,\nu} - q^2 A_\mu A_\nu \psi\right] + M^2\psi \pm \imath q\sqrt{E^2 - B^2 + 2\imath \vec{E}\cdot\vec{B}}\ \psi = 0 \tag{1}$$

In this, A_μ is the electromagnetic potential which generates fields $\vec{E}$ and $\vec{B}$, $q = e/(\hbar c)$, $M = (m_0 c)/\hbar$, e is the electronic charge, and m_0 is the electron rest mass. The actual scalar curvature of the geometry, B (not to be confused with the magnitude of the magnetic field vector $\vec{B}$), is given by

$$B = \psi^{-2} \tag{2}$$

and it is clearly complex valued.

The appendix in reference [1] demonstrates that solutions of equation (1) match solutions to Dirac's Equation for the case of uniform, constant, non-null electromagnetic fields. In that proof, the Dirac Equation itself is taken to be the (chiral) second order form of the equation,

$$\hat{\eta}^{\mu\nu}\left[\psi_{,\mu,\nu} + \imath q A_{\mu,\nu}\psi + 2\imath q A_\mu \psi_{,\nu} - q^2 A_\mu A_\nu \psi\right] + M^2\psi + \imath q\vec{\sigma}\cdot\left(\vec{E} + \imath\vec{B}\right)\psi = 0 \tag{3}$$

As a quick reference summary[1–3], equation (1) follows from a Weyl-like Cartan geometric model with gauge invariant variables defined for cases in which the curvature $B \neq 0$. For the metric, that definition is

$$\hat{g}_{\mu\nu} = (B/C)\, g_{\mu\nu} \tag{4}$$

where $C = \pm 1$. This is just the product of the scalar curvature and the gauge varying metric, but with the case $C = -1$ included originally to handle cases of $B < 0$, and retained as a legitimate flexibility of equation (4) even though $C = 1$ is used here[4–6].

---

*jrankin@panix.com

The analog to equation (4) for the Weyl vector is

$$\begin{aligned}\hat{v}_\mu &= v_\mu - [\tfrac{1}{2}\, ln\,(B/C)]_{,\mu} \\ &= v_\mu - (\tfrac{1}{2}\, ln\, B)_{,\mu}\end{aligned} \tag{5}$$

where in common electromagnetic gauge choices,

$$\begin{aligned}v_\mu &= \imath[e/(\hbar c)]A_\mu \\ &= \imath q A_\mu\end{aligned} \tag{6}$$

with A_μ real. Then

$$\begin{aligned}p_{\mu\nu} &= v_{\nu,\mu} - v_{\mu,\nu} \\ &= \hat{v}_{\nu,\mu} - \hat{v}_{\mu,\nu} \\ &= \hat{p}_{\mu\nu} \\ &= \imath q F_{\mu\nu}\end{aligned} \tag{7}$$

where $F_{\mu\nu}$ is the standard Maxwell tensor, the curl of A_μ, and it is real.

These relations together with the geometric kinematics imply that the gauge invariant variables are not independent, but must instead obey the kinematic identity

$$\hat{R} + 6\hat{v}^\mu{}_{\|\mu} + 6\hat{v}^\mu \hat{v}_\mu + \hat{a}^{\mu\nu}\hat{p}_{\mu\nu} = C(1 + \tfrac{1}{4}\,\hat{a}) \tag{8}$$

Here the "$\|$" derivative is the Riemannian covariant derivative based on $\hat{g}_{\mu\nu}$, $\hat{a}_{\mu\nu}$ is the self dual antisymmetric part of the metric, and $\hat{a} = \hat{a}_{\mu\nu}\hat{a}^{\mu\nu}$ (indices raised by $\hat{g}^{\mu\nu}$). This identity is the source of equation (1) once the dynamics is also specified, and that dynamics is given by the action integral

$$\begin{aligned}I = \int &[(\hat{R} - 2\sigma) - \tfrac{1}{2}\, j^2(\hat{p}_{\mu\nu}\hat{p}^{\mu\nu}) + \tfrac{1}{2}\{\hat{\beta}[\hat{R} + 6\hat{v}^\mu{}_{\|\mu} + 6\hat{v}^\mu\hat{v}_\mu + \hat{a}^{\mu\nu}\hat{p}_{\mu\nu} \\ &-C(1 + \tfrac{1}{4}\,\hat{a})] + 2\hat{\lambda}^{\mu\nu}\hat{p}_{\mu\nu} + \hat{\gamma}(\hat{a} - K^2) + CC\}]\sqrt{-\hat{g}}\, d^4x\end{aligned} \tag{9}$$

where the CC is the complex conjugate of all that precedes it in the brackets, K is a constant (equal to $6\imath$), $\hat{\lambda}^{\mu\nu}$ is real, thus constraining $\hat{p}_{\mu\nu}$ to be imaginary to fit equation (6), and the constraint $\hat{a} = K^2$ is simply a zero order approximation (based on geometric considerations) to some better theory of the antisymmetric part of the metric yet to be determined. Such a better theory should become a theory of rest mass, among other things.

In the sections to follow, the complex scalar nature of B will lead to the conclusion that at least in some sense, the full Dirac wavefunction ψ of equation (3) can be treated as a quaternionic spacetime scalar rather than necessarily always having spacetime spinor properties. It will be seen that this scalar property follows by projecting ("spin-up" and "spin-down") very general solutions to the equivalent quaternionic form of Dirac Equation (3) into the complex scalar solutions of equation (1) in the complex plane. This projecting preserves the full content of the Dirac ψ, with no information lost, and so it can be considered an embedding of Dirac Theory. Those sections will then be followed by a discussion of predicted quantum and other effects, and then finally the appendix will present an introduction to a generalized structure in which the scalar curvature B and gauge transformations may become general, quaternionic quantities.

II. QUATERNIONIC DIRAC THEORY

This section will present an immediate and simple translation of equation (3) into the equivalent, real quaternion form. To keep this as simple as possible, no effort will be made here to go into general formulations of quaternionic quantum mechanics[7, 8]. Rather, the two component chiral spinor will be translated into a single real quaternion variable obeying an immediately analogous equation to the spinor form, and still clearly containing the spinor predecessor.

A. Quaternions

In this presentation, the basic quaternions will be taken as abstract mathematical objects similar to the number 1 and the imaginary unit $\imath$ in complex numbers. Specifically, they are taken to be the four quantities Q_μ, where the

subscript does *not* imply the quantities are a four vector, and where

$$Q_0 = \sigma_0 \tag{10}$$

and

$$Q_k = -\imath\sigma_k \tag{11}$$

for $k = 1, 2, 3$, where the σ_k are the standard Pauli spin matrices[9], and σ_0 is the unit matrix. The basic properties of quaternions are reviewed in many references, such as Adler[8], and Morse and Feshbach[10], and there are many representations of them which may differ from those of equations (10) and (11), yet which are algebraically isomorphic to those quantities. Such possibly isomorphic representations will be denoted here by Q'_μ, and they may happen to be assigned particular coordinate transformation properties for convenience, unlike the Q_μ, which are mathematical invariants.

The Q_k have an obvious vector form

$$\vec{Q} = \sum_{k=1}^{3} \hat{e}_k Q_k \tag{12}$$

where the $\hat{e}_k$ are the Cartesian unit vectors. A completely analogous form exists for the Q'_k,

$$\vec{Q}' = \sum_{k=1}^{3} \hat{e}'_k Q'_k \tag{13}$$

although now the $\hat{e}'_k$ may be unit vectors in one of the curvilinear coordinate systems in common use[10].

For definiteness, choose a set of $Q'_\mu = Q_\mu$. A general real quaternion variable will be any quantity of the form

$$\psi = \psi_{0R}Q_0 - \psi_{1I}Q_1 + \psi_{1R}Q_2 - \psi_{0I}Q_3 \tag{14}$$

where the coefficients ψ_{nR} and ψ_{nI} are all real numbers. The notation has been chosen with an eye to the wavefunction values, and in matrix form it is

$$\psi = \begin{pmatrix} \psi_{0R} + \imath\psi_{0I} & -\psi_{1R} + \imath\psi_{1I} \\ \psi_{1R} + \imath\psi_{1I} & \psi_{0R} - \imath\psi_{0I} \end{pmatrix} \tag{15}$$

This is in fact the quaternion representation of the spinor wavefunction

$$\psi = \begin{pmatrix} \psi_{0R} + \imath\psi_{0I} \\ \psi_{1R} + \imath\psi_{1I} \end{pmatrix} \tag{16}$$

where the notation ψ_{nR} and ψ_{nI} now clearly gives the real and imaginary parts of the two rows in the spinor.

Notice that the quaternion has no additional information over the spinor, because the second column only contains conjugate forms of quantities in the first column. The exact relationship of the columns, given in equation (15), is common to all real quaternions, including real quaternion equations.

B. Translation from Spinor Equation to Quaternion Equation

Equations (16) and (15) already give the translation of ψ from spinor to quaternion. A spinor equation such as equation (3) will contain the spinor ψ with various operations on it, including products of it with coefficients which may be complex, and are assumed to commute with it, and also the product of $\vec{\sigma}\cdot\left(\vec{E} + \imath\vec{B}\right)$ with it. The factor $\vec{\sigma}$ in the spin term is *not* assumed to commute with ψ.

Now in order to handle the factor $\imath$ in a coefficient, note that for the quaternion ψ,

$$-\psi Q_3 = \begin{pmatrix} \imath(\psi_{0R} + \imath\psi_{0I}) & -\imath(-\psi_{1R} + \imath\psi_{1I}) \\ \imath(\psi_{1R} + \imath\psi_{1I}) & -\imath(\psi_{0R} - \imath\psi_{0I}) \end{pmatrix} \tag{17}$$

But the first column of this is clearly $\imath$ times the original column, while the second column maintains the correct conjugate relationships for a quaternion. In point of fact, the choice of $-Q_3$ can be generalized to other fixed

directions in $\hat{n}\cdot\vec{Q}$, where $\hat{n}$ is an arbitrary constant unit vector. However this will not be done here in order not to lose immediate transparency between quaternion and spinor forms.

Equation (17) leads to a general rule for translating a factor $\imath$ multiplying a spinor into the equivalent quaternion value, and that is

$$\imath \to |(-Q_3) \tag{18}$$

In this, the "|" is the "leap-over operator" or "barred operator", which signifies that the quantity immediately following it is to be applied to the right side of any product of further quantities following the term in a product. This operation is required to handle the possible two-sided nature of any quaternion product (a "split product"), and has been in use in the literature for some time by researchers of quaternionic Dirac Theory, such as De Leo and Rotelli[11], and Schwartz[12]. For example, the spinor eigenvalue equation related to the z component of angular momentum translates via

$$\psi_{,3} = \imath m\psi \to m|(-Q_3)\psi = -m\psi Q_3 \tag{19}$$

For completeness, the leap-over operation "$(X)\|$" is also defined, and will denote that X is applied to the left side of the product of any quantities preceding it in a product. In practice, the actual position in a leap-over operation is determined when the equation is evaluated explicitly or implicitly (for ψ in this case, which is normally considered free of leap-over operators, or at worst, to contain only $|(-Q_3)$). If the position is then shifted somehow without reevaluating the equation, contradictory, invalid results can appear.

It will be noted that $|(-Q_3)$ does in fact effectively commute with a standard quaternion value ψ by definition, and so it has the same property that $\imath$ had in the spinor equation vis-à-vis the spinor ψ. This property is actually quite useful in handling spinor/quaternion expression translations back and forth. Of course, the translation back to spinor form involves merely using the first column of the quaternion expression. The spinors remain explicitly visible in the first column throughout the quaternion expression.

This still leaves the product of ψ with $\vec{\sigma}\cdot\left(\vec{E}+\imath\vec{B}\right)$ in the spin term, in which the $\vec{\sigma}$ will not commute with the spinor ψ. But equations (11) and (12) imply that

$$\vec{\sigma} = \imath\vec{Q} \tag{20}$$

This and equation (18) then give immediately the full real quaternion translation of spinor equation (3) as

$$\begin{aligned}\hat{\eta}^{\mu\nu}\left[\psi_{,\mu,\nu} - qA_{\mu,\nu}\psi Q_3 - 2qA_\mu\psi_{,\nu}Q_3 - q^2A_\mu A_\nu\psi\right]&\\ +M^2\psi - q\vec{Q}\cdot\left(\vec{E}\psi - \vec{B}\psi Q_3\right) = 0&\end{aligned} \tag{21}$$

This could also be written as

$$\begin{aligned}\hat{\eta}^{\mu\nu}\left[\psi_{,\mu,\nu} - qA_{\mu,\nu}\psi Q_3 - 2qA_\mu\psi_{,\nu}Q_3 - q^2A_\mu A_\nu\psi\right]&\\ +M^2\psi - q\vec{Q}\cdot\left(\vec{E} - \vec{B}|Q_3\right)\psi = 0&\end{aligned} \tag{22}$$

and clearly other alternative forms using $|(-Q3)$ are possible. It is also possible to use $\vec{Q}'$ and curvilinear coordinate forms in the spin term, and later it will be clear that various "spin space" transformations will allow passage between different representations of the $\vec{Q}'$. When transforming the $\vec{Q}'$, the $|(-Q3)$ quantities all remain unprimed and unchanged, because that term is treated as a coordinate invariant, like $\imath$ in the coefficients in the spinor forms.

It is also possible to translate the complex equation (1) into quaternion form for comparison. It becomes

$$\begin{aligned}\hat{\eta}^{\mu\nu}\left[\psi_{,\mu,\nu} - qA_{\mu,\nu}\psi Q_3 - 2qA_\mu\psi_{,\nu}Q_3 - q^2A_\mu A_\nu\psi\right]&\\ +M^2\psi \pm q\sqrt{-\left[E^2 - B^2 - 2\vec{E}\cdot\vec{B}|Q_3\right]}\;\psi = 0&\end{aligned} \tag{23}$$

where the quaternion form of ψ here is

$$\psi = \begin{pmatrix} \psi_R + \imath\psi_I & 0 \\ 0 & \psi_R - \imath\psi_I \end{pmatrix} \tag{24}$$

or equivalently,

$$\psi = \psi_R Q_0 - \psi_I Q_3 \tag{25}$$

This is how a complex number appears in this quaternion form.

A comparison of equations (22) and (23) now suggests immediately that in many ways they are *the same* equation. This is seen by noting that

$$\left[-\vec{Q}'\cdot\left(\vec{E}-\vec{B}|Q_3\right)\right]\left[-\vec{Q}'\cdot\left(\vec{E}-\vec{B}|Q_3\right)\right]= \\ -\left[E^2-B^2-2\vec{E}\cdot\vec{B}|Q_3\right] \tag{26}$$

Clearly both equations have the same general form in which the spin terms are the square root of the same quantity. However, they are different roots from the quaternion viewpoint.

Equation (23) extracts the roots within the complex plane, while equation (22) uses a root that extends into more general quaternion space. Nevertheless, the solutions of equation (22) when the electromagnetic field is non-null can be expanded in the eigenspinors/eigenquaternions of the electromagnetic field. The "spin-up" and "spin-down" projections of the result will each separately obey a form of equation (23) with a superposition of terms with both plus and minus values of the square root in the equation. This is not an eigenstate of equation (23) generally, which would have to contain only a single sign for the square root term. However, generally eigenstates of equation (23) do not have real energy eigenvalues. Rather, those energy eigenvalues are complex valued, with corresponding poor behavior in the eigenfunctions, although exceptions include the case of uniform, constant, non-null, electromagnetic fields[1]. However, the energy eigenvalues of equation (3) (or equation (21)) are well known to be satisfactory with well-behaved eigenfunctions. The "spin-up" and "spin-down" projections of those solutions inherit the eigenvalues and that good behavior. Thus, the Dirac Equation provides well-behaved solutions to equation (23), a subject that will be addressed more thoroughly in the next section. However, some additional points about Dirac Theory and quaternionic Dirac Theory deserve mention first.

C. Quaternion Form of Dirac Theory

The translation from spinor to quaternion can be extended to first order Dirac Theory and a Lagrangian[11], and a full formalism developed. The Dirac γ^μ translate just like other expressions in spinor equations. For example, the chiral representation giving equation (22) corresponds to

$$\gamma^0=\begin{pmatrix} 0 & -\imath\sigma_0 \\ -\imath\sigma_0 & 0\end{pmatrix}\rightarrow\begin{pmatrix} 0 & Q_0'|Q_3 \\ Q_0'|Q_3 & 0\end{pmatrix} \tag{27}$$

and

$$\gamma^k=\begin{pmatrix} 0 & \imath\sigma_k \\ -\imath\sigma_k & 0\end{pmatrix}\rightarrow\begin{pmatrix} 0 & -Q_k' \\ Q_k' & 0\end{pmatrix} \tag{28}$$

for $k=1,2,3$, in the first order Dirac Equation

$$\gamma^\mu\left(\zeta_{,\mu}-qA_\mu\zeta Q_3\right)=M\zeta \tag{29}$$

where ζ is a column vector containing two rows of quaternions. Clearly "spin space" transformations in Dirac Theory become transformations now on the Q_μ', the internal quaternion basis vectors. Such transformations will transform between various representations of the Q_μ'.

The particular form of γ^0 causes the formation of ${}^\dagger\gamma^0$ (the spinor or quaternion conjugate, with the "$\dagger$" written to the left of its target γ^0 rather than to the right) to present more than one suggested analog translating from spinor to quaternion form, depending on whether the "$\dagger$" is taken before, or after, translating from spinor to quaternion. The two possibilities are always related to each other in that either one can be obtained from the other by multiplying it by $[-(Q_3)\|\ |(Q_3)]$. However, generally most of spinor Dirac Theory translates fairly clearly, and that will not be detailed here, but there are a few points relating to the theory which produced equation (1) originally which will be mentioned.

From the viewpoint of the theory of equation (1), second order Dirac Theory is expected to be more fundamental than first order Dirac Theory, which can be derived from a single, second order, chiral equation[13]. In that same spirit, there is a spinor action principle which leads directly to the second order, spinor Dirac Equation forms, as opposed to the standard Dirac action which leads to the first order equations[10, 11, 14]. The second order equations follow from

$$\begin{aligned} I \; = \; & \int \{[\hbar/(4M)][-\imath\eta^{\mu\nu}(\bar{\zeta}_{,\mu} - \imath q A_\mu \bar{\zeta})(\zeta_{,\nu} + \imath q A_\nu \zeta) + \imath M^2 \bar{\zeta}\zeta \\ & -\imath \bar{\zeta}_{,\mu}\Sigma^{\mu\nu}\zeta_{,\nu} - q A_\mu (\bar{\zeta}\Sigma^{\mu\nu}\zeta)_{,\nu} + CC] - [1/(16\pi c)]F_{\mu\nu}F^{\mu\nu}\} d^4x \end{aligned} \tag{30}$$

where the CC is the complex conjugate of all that precedes it in the brackets, $\bar{\zeta} = {}^\dagger\zeta\gamma^0$, and

$$\Sigma^{\mu\nu} = -\frac{1}{2}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu) \tag{31}$$

However, the action of equation (30) does *not* give standard Dirac Theory. This is easily seen in a chiral representation, where the two chiral parts of ζ will each obey a second order Dirac Equation form, but without any first order Dirac Equations relating them to each other. The two chiral parts are actually independent of each other. J. T. Wheeler pointed out to me that this independence is equivalent to the presence of a negative energy density ghost wavefunction field, in addition to the standard positive energy density wavefunction field[2, 15]. Moreover, the same situation arises in the theory that leads to equation (1), meaning that it also includes such a ghost. For that reason, the action of equation (30) remains of interest here, since it generates the analogous Dirac Theory generalization. And, it is also more amenable to modifications which yield the general relativistic "conformal term" inherited from equation (8) instead of the $\frac{1}{4}\hat{R}\psi$ type of term that results in standard Dirac Theory from the interaction of spin and the Riemann Tensor (see the appendix in reference [1]). Clearly equation (8) points to a coefficient of $\frac{1}{6}$ rather than $\frac{1}{4}$. That point is not trivial, because if confirmed by observation, it would imply that the standard, first order Dirac Equations are only valid in the special relativistic limit. As they currently stand, when General Relativity is important, they would not give the correct conformal term in the second order form of the wave equation.

Besides the antisymmetric part of the metric, $\hat{a}_{\mu\nu}$, this ghost, negative energy density "Wheeler" field, is an additional classical field that appears in this framework. And, since it is a *classical* framework, it is worth noting that classical physics has previously shown tolerance for a negative energy density field in its midst. Prior to the appearance of General Relativity, gravitation was described by the Newtonian Gravitational Field[16], and that field is necessarily a negative energy density field. Work is extracted from the field as the field is built up, so it is a negative energy density field by definition, yet classical physics coexisted with it without catastrophe. For this reason, I take the ghost field as seriously worth examining. Indeed, a similar pair of positive-negative energy density fields has been studied in some detail by Doroshkevich et al. in recent literature[17]. There will be a few more comments on this ghost, negative energy density field later.

III. EIGENQUATERNIONS AND COMPLEX/SPIN PROJECTIONS

A. Eigenspinors/Eigenquaternions

Consider the eigenspinor equation

$$\imath\vec{\sigma}\cdot\left(\vec{E} + \imath\vec{B}\right)\xi = -\vec{Q}'\cdot\left(\vec{E} + \imath\vec{B}\right)\xi = \lambda\xi \tag{32}$$

and its quaternion equivalent

$$-\vec{Q}'\cdot\left(\vec{E} - \vec{B}|Q_3\right)\xi = \lambda\xi \tag{33}$$

In these two equations, ξ is a spinor and its equivalent quaternion respectively, and λ is an eigenvalue which commutes with all the other terms in both cases. These ξ are essentially the eigenspinors/eigenquaternions of the electromagnetic field[1, 18]. Hereafter, the term "eigenvector" will be used to denote these quantities which can be expressed as either spinors or quaternions.

Since λ commutes with everything in the equation, the operator $-\vec{Q}'\cdot\left(\vec{E} - \vec{B}|Q_3\right)$ can be applied to ξ twice. Then equation (26) gives that

$$\lambda^2 = -\left[E^2 - B^2 - 2\vec{E}\cdot\vec{B}|Q_3\right] \tag{34}$$

For non-null fields, this has the two distinct values

$$\lambda_{\pm} = \pm\sqrt{-\left[E^2 - B^2 - 2\vec{E}\cdot\vec{B}|Q_3\right]} \tag{35}$$

where the square root is understood to be taken as a complex number of the general form $\lambda_R + [|(-Q_3)]\,\lambda_I$ with λ_R and λ_I real. But these two values are exactly the two choices for the root in equation (23). Furthermore, the values of λ do commute with everything in equation (33) as required, because $|(-Q_3)$ effectively commutes with all the other terms. This commutation includes the fact that $|(-Q_3)$ clearly commutes with $|(-Q_3)$, and additionally, $|(-Q_3)\,|(-Q_3) = -1$.

For each of the two values of λ, the matching ξ is the eigenvector for that eigenvalue, either ξ_+ or ξ_-. Since the fields are assumed to be non-null, the eigenvalues are nonzero and different, and the two eigenvectors are independent[1, 18]. Thus, an arbitrary spinor/quaternion ψ can be expanded in terms of the two as

$$\psi = a\xi_+ + b\xi_- \tag{36}$$

This is most easily done with the eigenspinors, with translation to eigenquaternions following the expansion. Then the coefficients a and b are complex functions just as λ is, and they commute with everything at least as well as λ does in both forms.

Now define

$$\psi_+ = a\xi_+ \tag{37}$$

and

$$\psi_- = b\xi_- \tag{38}$$

so that $\psi = \psi_+ + \psi_-$. But now, equation (33) implies that

$$-\vec{Q}'\cdot\left(\vec{E} - \vec{B}|Q_3\right)\psi = \lambda_+\psi_+ + \lambda_-\psi_- \tag{39}$$

and equation (22) becomes

$$\begin{aligned}\hat{\eta}^{\mu\nu}[(\psi_+ \;&+\; \psi_-)_{,\mu,\nu} - qA_{\mu,\nu}(\psi_+ + \psi_-)Q_3 \\ &-\; 2qA_\mu(\psi_+ + \psi_-)_{,\nu}Q_3 - q^2A_\mu A_\nu(\psi_+ + \psi_-)] \\ &+\; M^2(\psi_+ + \psi_-) + q\lambda_+\psi_+ + q\lambda_-\psi_- = 0\end{aligned} \tag{40}$$

This is now a superposition of forms similar to equation (23), but the $\psi_\pm$ are not actually complex numbers at this point. Rather they are still more general quantities like the Dirac ψ. If the electromagnetic field is also constant and uniform, the eigenvectors can be chosen as constant vectors, and dotted into equation (40) to obtain an equation for either coefficient a or b of the form of equation (23) with a single choice of the sign of the square root, a pure eigenstate of that equation[1]. But for more general electromagnetic fields, a different technique must be used to obtain the form of equation (23). The more general result will give a superposition of the forms from that equation for both values of the square root, rather than a single eigenstate.

B. Complex/Spin Projections

The standard projection operations in quantum mechanics for “spin-up” and “spin-down” are[19]

$$\psi_u = \frac{1}{2}[\psi + \sigma_z\psi] \tag{41}$$

and

$$\psi'_d = \frac{1}{2}[\psi - \sigma_z\psi] \tag{42}$$

where the next steps will reveal why the second quantity is primed.

These have immediate quaternionic translations, with "spin-up" becoming

$$\psi_u = \frac{1}{2}[\psi - Q_3 \psi Q_3] \tag{43}$$

This is immediately recognizable as the complex projection of the quaternionic ψ into the complex plane defined by quantities having the form of equation (25)[11]. The translation of equation (42) can be put in this same form simply by multiplying after translation from the left by $-Q_2$, giving

$$\psi_d = \frac{1}{2}[(-Q_2)(\psi + Q_3 \psi Q_3)] \tag{44}$$

The reason for the prime in equation (42) is now clear, since a full equivalence to equation (44) would require ψ'_d to have its nonzero component shifted to the top position by prefixing the right side of equation (42) with a multiplier of $\imath\sigma_y$ to match the $-Q_2$ in equation (44).

In matrix form, using equation (15) as ψ, equations (43) and (44) are

$$\psi_u = \begin{pmatrix} \psi_{0R} + \imath\psi_{0I} & 0 \\ 0 & \psi_{0R} - \imath\psi_{0I} \end{pmatrix} \tag{45}$$

and

$$\psi_d = \begin{pmatrix} \psi_{1R} + \imath\psi_{1I} & 0 \\ 0 & \psi_{1R} - \imath\psi_{1I} \end{pmatrix} \tag{46}$$

and both are clearly complex in the sense of equations (24) and (25). Furthermore, the original quaternionic ψ can be fully reconstructed from this pair as

$$\psi = \psi_u + Q_2 \psi_d \tag{47}$$

so no information is lost in these projections.

Now construct ψ_{+u}, ψ_{+d}, ψ_{-u}, and ψ_{-d} using equations (43) and (44) as a pattern. Since Q_2 (from the left) and Q_3 commute with the derivatives and the coefficients of every term in equation (40), it is immediately clear that one can add and subtract the results of the various operations with $-Q_2$ and Q_3 on that equation to get

$$\begin{aligned} \hat{\eta}^{\mu\nu}[(\psi_{+u} &+ \psi_{-u})_{,\mu,\nu} - qA_{\mu,\nu}(\psi_{+u} + \psi_{-u})Q_3 \\ &- 2qA_\mu(\psi_{+u} + \psi_{-u})_{,\nu}Q_3 - q^2 A_\mu A_\nu(\psi_{+u} + \psi_{-u})] \\ &+ M^2(\psi_{+u} + \psi_{-u}) + q\lambda_+\psi_{+u} + q\lambda_-\psi_{-u} = 0 \end{aligned} \tag{48}$$

and

$$\begin{aligned} \hat{\eta}^{\mu\nu}[(\psi_{+d} &+ \psi_{-d})_{,\mu,\nu} - qA_{\mu,\nu}(\psi_{+d} + \psi_{-d})Q_3 \\ &- 2qA_\mu(\psi_{+d} + \psi_{-d})_{,\nu}Q_3 - q^2 A_\mu A_\nu(\psi_{+d} + \psi_{-d})] \\ &+ M^2(\psi_{+d} + \psi_{-d}) + q\lambda_+\psi_{+d} + q\lambda_-\psi_{-d} = 0 \end{aligned} \tag{49}$$

These clearly give us two separate superpositions of the form of equation (23) as solutions. Furthermore, the Dirac eigenvalues propagate intact into these solutions, so that the solutions do have good eigenvalues and well behaved functions. On the other hand, these do not represent a sum of *eigenfunctions* of equation (23), because such eigenfunctions are generally not as well behaved as the two parts of these two results.

These two solutions correspond to actual scalar curvatures B as given through equation (2) as

$$B_u = \psi_u^{-2} \tag{50}$$

and

$$B_d = \psi_d^{-2} \tag{51}$$

Clearly there is a "spin-up" value of B, and a separate "spin-down" value of B, both complex scalars. They hold all the information of the Dirac ψ between them. In that sense then, the full Dirac wavefunction ψ has been embedded into the basic geometric structure in the complex plane, but ψ viewed as a quaternionic spacetime scalar, not as a spinor. The fact that it has been so embedded suggests that the Dirac Equation has as much in common with the

complex plane as it has with quaternions. To add additional perspective on this point, the appendix to this paper discusses generalizing gauge transformations and curvatures in this structure to quaternionic values.

One more point deserves mention here. Both equation (1) and equation (3) are paired with natural conjugate wavefunction equations in their respective theoretical frameworks[1, 13]. Equation (1) pairs with

$$\hat{\eta}^{\mu\nu}\left[\xi_{,\mu,\nu} - \imath q A_{\mu,\nu}\xi - 2\imath q A_\mu \xi_{,\nu} - q^2 A_\mu A_\nu \xi\right] \\ + M^2\xi \pm \imath q\sqrt{E^2 - B^2 + 2\imath \vec{E}\cdot\vec{B}}\ \xi = 0 \tag{52}$$

The complex conjugate of this gives

$$\hat{\eta}^{\mu\nu}\left[{}^{\#}\xi_{,\mu,\nu} + \imath q A_{\mu,\nu}{}^{\#}\xi + 2\imath q A_\mu {}^{\#}\xi_{,\nu} - q^2 A_\mu A_\nu {}^{\#}\xi\right] \\ + M^2\,{}^{\#}\xi \mp \imath q\sqrt{E^2 - B^2 - 2\imath \vec{E}\cdot\vec{B}}\ {}^{\#}\xi = 0 \tag{53}$$

where a " $\#$ " prefixing a quantity designates the complex conjugate of that quantity. On the other hand, the conjugate Dirac Equation form paired with equation (3) is

$$\hat{\eta}^{\mu\nu}\left[\phi_{,\mu,\nu} + \imath q A_{\mu,\nu}\phi + 2\imath q A_\mu \phi_{,\nu} - q^2 A_\mu A_\nu \phi\right] \\ + M^2\phi - \imath q\vec{\sigma}\cdot\left(\vec{E} - \imath\vec{B}\right)\phi = 0 \tag{54}$$

Clearly one can develop a program of using the "spin up/down" projections of ϕ just as was done above with ψ. The result gives the superposition of forms of equation (53) for ${}^{\#}\xi$, thereby giving a solution for the conjugate ξ used in reference [1]. If the (chiral) Dirac ψ and ϕ are also related via the first order Dirac Equations[13], then the matching family of projected scalar functions ψ and ξ might be expected to have no more symptoms of negative energy density than the original Dirac solution has. Of course, the action of equation (30) does not require that the first order equations relate the chiral Dirac wavefunctions ψ and ϕ, as noted earlier. From its viewpoint, the first order equations are additional conditions.

IV. QUANTUM AND OTHER EFFECTS

The previous sections have demonstrated that Dirac Theory *does* project into the complex forms of the classical gauge theory in reference [1]. What then is predicted?

The most immediate prediction has already been mentioned, that the Dirac eigenvalues, as well as the projections of the eigenfunctions, propagate into the complex theory. That brings those quantities in the complex theory into line with Dirac, and it brings a form of spin 1/2 itself into the results. But what else does the formalism imply?

A. Effects and Scale of Nonnegligible Ricci Tensor

First note that there are "conformal terms" in the stress tensor produced by the wavefunctions in $\hat{\beta}$, where $\hat{\beta} = \xi\psi$, and ξ is the conjugate wavefunction in the theory derived from the action of equation (9)[1]. Those terms have the form $(\hat{\beta})_{\|\mu\|\nu} - (\hat{\beta})_{\|\gamma\|\tau}\hat{g}^{\gamma\tau}\hat{g}_{\mu\nu}$, and as noted in references [3] and [2], they originate from the $\hat{\beta}\hat{R}$ term in the action, or the equivalent, well-known conformal term in theories of scalar wavefunction fields[20]. However, their magnitude should be comparable to the "ordinary" terms in the stress tensor, since there is no higher order, multiplying factor in them. But the covariant divergence of these terms is

$$\left[(\hat{\beta})_{\|\mu}{}^{\|\nu} - (\hat{\beta})_{\|\gamma}{}^{\|\gamma}\delta^\nu_\mu\right]_{\|\nu} = -\hat{R}_\mu{}^\nu(\hat{\beta})_{,\nu} \tag{55}$$

This is clearly a general relativistic quantity, which means that in regions in which $\hat{R}_\mu{}^\nu \approx 0$, these terms have vanishing covariant divergence. In that case, they cannot interact with the rest of the stress tensor except gravitationally, no matter how much energy they contain. Otherwise, they are separately conserved and invisible outside regions in which $\hat{R}_\mu{}^\nu \neq 0$. Furthermore, they stem from wavefunctions now seen to be spin 1/2 wavefunctions. In some sense, they represent nearly noninteracting, spin 1/2 matter-energy. The *effective* rest mass of such terms may only be clear after integrating their stress tensor terms to get a momentum four vector, and contracting that with itself.

As previously stated, these terms have long been present as "conformal terms" in theories of scalar wavefunction fields such as the Klein-Gordon field[2], where the analogous terms are those in which $\hat{\beta} \to {}^{\#}\psi\psi$ in the above

expressions. One simply modifies the rest mass squared term in the standard action for such fields by the substitution $M^2 \to M^2 + \frac{1}{6}R$, where R is the scalar curvature. However, it is fairly easy to demonstrate that the conformal stress tensor terms must vanish whenever the ordinary stress tensor terms of a Klein-Gordon wavefunction field vanish, because the ordinary energy density is a sum of positive definite quantities unless $\psi = 0$. But $\psi = 0$ implies $\left({}^{\#}\psi\psi\right)_{\|\mu\|\nu} - \left({}^{\#}\psi\psi\right)_{\|\gamma\|\tau} \hat{g}^{\gamma\tau}\hat{g}_{\mu\nu}$ is also zero, and the conformal terms then really completely vanish also. In the conformal Klein-Gordon case, they are rather tightly linked to the presence of "ordinary matter", and have spin 0 as well, not spin 1/2.

In that regard, the formalism of this paper has already been noted to be spin 1/2, and because of the possibility of negative energy density, its energy density is no longer necessarily positive definite. It's not clear at this point if that allows the conformal stress energy terms to be nonzero while the ordinary stress energy terms vanish. However, that's the basic requirement to delink the two sets of terms from each other significantly.

Equation (55) indicates that interaction between the conformal terms and the rest of the stress tensor *does* occur in regions in which $\left|\hat{R}_{\mu}{}^{\nu}\right| \geq 1$. Using the Reissner-Nordström solution of an electronic charge[21] as a guide to calculate $\hat{R}_{\mu\nu}$[2],

$$\left|\hat{R}_{\mu\nu}\right| \sim \frac{3}{4}[(Gm_0^2)/(\hbar c)][e^2/(\hbar c)](1/r^4) \tag{56}$$

In this, the factors of m_0, e, and $\hbar$ are introduced through the scale factor b_0 as well as the coefficient of A_μ in equation (6), and the scale factor is set to match the wave equation to atomic scale phenomena[1]. The magnitude of $\hat{R}_{\mu\nu}$ approaches unity for a dimensionless radius $r_0 = 1.8 \cdot 10^{-12}$. Convert that to lab units by dividing the dimensionless value by the square root of the scale factor

$$b_0 = \frac{3}{4}\left[\left(m_0^2c^2\right)/\hbar^2\right] \tag{57}$$

and the result gives $r_{0\,LAB} = r_0/\sqrt{b_0} \approx 5 \cdot 10^{-22}$ *cm.* (a much larger value than the 10^{-34} *cm.* at which the metric deviates significantly from flat spacetime values). This represents a slight shift from values calculated in reference [2], partly because of a more accurate treatment of the Reissner-Nordström solution, and partly because a different universal scale factor b_0 is indicated by the more general formalism which includes an antisymmetric part to the metric[1].

As noted in references [2] and [3], any corresponding interaction cross section should contain the factor $\pi r_{0\,LAB}^2$, so this crude estimate gives a very approximate cross section of about 10^{-42} *cm.*^{2}. This is just above the lower bound of observed neutrino cross sections in matter[22]. However, this isn't just the cross section at which the conformal terms interact with the rest of the stress tensor. It's also the cross section at which *any* general relativistic effects of the Reissner-Nordström solution that depend upon nonzero $\hat{R}_{\mu\nu}$ could become important. Again, these general relativistic effects are caused by the electric charge term in the Reissner-Nordström metric, which produces these effects in $\hat{R}_{\mu\nu}$ at a much larger radius than either its effects on the metric, or the metric effects of the Schwarzchild type mass term in the solution[21]. That mass term does not contribute to a nonzero $\hat{R}_{\mu\nu}$ anyway.

B. Scale of Nonnegligible Scalar Curvature

There are also non-electromagnetic contributions to $\hat{R}_{\mu\nu}$ as well as $\hat{R}$ from the $\hat{\beta}$ terms in the stress tensor, but they occur at magnitudes beneath the above effects. To get a feel for their magnitude, directly compare the forms generated for the "rest mass term" in the stress tensors from the action of equations (9), and the Klein-Gordon action in the appendix in reference [2]. The Klein-Gordon action can be used for this purpose since the rest mass term in it is the same magnitude as the analogous term in the action of equation (30), and the scalar nature of the earlier ψ is easier to handle. However, the Klein-Gordon case should be rendered dimensionless by multiplying the stress tensor by $(1/b_0)$ to allow direct comparison of dimensionless quantities. That $(1/b_0)$ multiplier is absorbed into the derivatives involved in the calculation of curvature so that the derivatives are consistently taken with respect to the same dimensionless variables as are used in the wave equation (8), and throughout the formalism.

The pertinent terms are $-\frac{3}{2}\left[\frac{1}{6}(1+\frac{1}{4}\hat{a})\hat{\beta} + CC\right]$ and $-[(k\hbar c)/(2b_0)]M_A\,{}^{\#}\psi\psi$ respectively, where $M_A = (m_A c)/\hbar$, m_A is the actual rest mass of the system in question, and $k = (8\pi G)/c^4$. But also, the identification is made that $\frac{1}{6}(1+\frac{1}{4}\hat{a}) = -M_A^2/b_0$, and positive energy density wavefunctions ζ_2 are chosen in $\hat{\beta}$. Reference [2] details the introduction of ζ_2, but in brief, $\hat{\beta} = \xi\psi$, and ζ_1 and ζ_2 are defined by

$$\psi = (1/\sqrt{2})(\zeta_1 + \zeta_2) \tag{58}$$

and

$$\xi = (1/\sqrt{2}\,)(\,^{\#}\zeta_1 - \,^{\#}\zeta_2) \tag{59}$$

This splits results into positive and negative energy density fields[15], and the negative energy density in ζ_1 is discarded for this calculation, giving $(\hat{\beta} + CC) = -^{\#}\zeta_2\zeta_2$. All these together then give

$$^{\#}\zeta_2\zeta_2 = (8\pi/3)[(\hbar G)/c^3][\hbar/(m_A c)]\,^{\#}\psi\psi \tag{60}$$

where the factor $(\hbar G)/c^3 = \mathcal{L}_P^2$, and $\mathcal{L}_P$ is the well-known Planck length. Since ψ is the standard laboratory quantum mechanical wavefunction, $^{\#}\psi\psi$ integrates to unity over its containing volume, giving $^{\#}\psi\psi$ the dimensions of inverse volume. This is consistent with $^{\#}\zeta_2\zeta_2$ (and $\hat{\beta}$) being dimensionless, as required. But this relation is all that is needed to relate the magnitudes of ψ and ζ_2 via

$$\zeta_2 = \{(8\pi/3)[(\hbar G)/c^3][\hbar/(m_A c)]\}^{1/2}\psi \tag{61}$$

Now ignoring the cosmological constant in reference [1], the expression for $\hat{R}$ is

$$\begin{aligned} \hat{R} &= \frac{1}{2}\left[(1 + \tfrac{1}{4}\,\hat{a})\hat{\beta} + CC\right] \\ &= 3(M_A^2/b_0)\,^{\#}\zeta_2\zeta_2 \\ &= 8\pi(M_A^2/b_0)[(\hbar G)/c^3][\hbar/(m_A c)]\,^{\#}\psi\psi \\ &= 8\pi(M_A/b_0)[(\hbar G)/c^3]\,^{\#}\psi\psi \end{aligned} \tag{62}$$

Unlike $\hat{R}_{\mu\nu}$, $\hat{R}$ contains no electromagnetic contribution. Thus it will be terms containing $\hat{R}$ that are sensitive to $(\hat{\beta} + CC) = -^{\#}\zeta_2\zeta_2$ approaching unit magnitude, such as the conformal term in the wave equation itself. A simple calculation reveals that an electron would have to be confined tightly within a sphere with a radius of roughly $10^{-25}cm$. to have a wavefunction magnitude great enough to produce an $\hat{R}$ with a magnitude approaching unity. This radius is three orders of magnitude beneath the radius at which $\hat{R}_{\mu\nu}$ approaches unit magnitude in the Reissner-Nordström solution, as discussed above. It is still about eight orders of magnitude above a Planck scale radius.

C. Quantized Energy Exchange Between Fields

What energy values are given when the $\hat{\beta}$ terms in the stress tensor generated by equation (9) are actually integrated? I investigated this for the hydrogen atom field shortly after publication of reference [1] in hopes that the results would agree better with Dirac Theory than the complex energy eigenvalues of equation (1). Using an adaptation of Lin's integrals as explained by Infeld and Hull[23] – and with help from some simplifying assumptions before I understood that energy density could become negative – I found that to lowest order, integrated energy values agreed with the energy eigenvalues of nonrelativistic Schrödinger Theory[9]. It's to this same order that the eigenvalues of equation (1) are real, and the eigenfunctions well behaved, so wavefunction energy eigenvalues tracked in tandem with the integrated energy values to the same degree that the results were well behaved overall. In retrospect, this should not be surprising, since had this not been true for the similar, original field theories of quantum wavefunctions[10], those theories would have been internally inconsistent with quantized atomic energies. Nevertheless, this has an important implication.

Ignoring general relativistic effects (gravitation), the total stress tensor energy is conserved. Since the wavefunction portion carries the discrete energy amounts predicted by wave mechanics, it is necessarily true that the electromagnetic field *must* exchange energy discretely with the bound state wavefunctions. But that is a major part of Planck's original quantum postulate[9]. All that remains is to demonstrate that the electromagnetic radiation emitted or absorbed has frequency $\nu = E/h$ where E is the discrete energy amount transferred, and the full Planck postulate would be respected. That would account correctly for blackbody radiation.

Lamb and Scully demonstrate that the photoelectric effect can be explained without resort to formal quantum field theory[24], while Thorn et al.[25] discuss just what experimental issues really do need to be addressed to measure genuine quantum field theoretic phenomena. Lamb and Scully list four phenomena to attribute to quantum field theory, including spontaneous emission, blackbody radiation, the Compton Effect, and electrodynamic level shifts. At this time, it appears possible that at least one of these might be explained by the approach of this paper, and that at least some form of electromagnetic quanta exist in this theory. This needs further examination.

V. CONCLUSIONS

The second order, chiral form of the Dirac Equation in flat spacetime does indeed "spin up/down" project neatly into the complex plane as a superposition of forms predicted by the classical gauge theory of reference [1]. That theory thus did include Dirac solution forms and spin 1/2 generally, *not* just in the special case of uniform, constant, non-null fields. The Dirac wavefunction field is seen to project into scalars in the complex plane, and in this sense, it can be viewed naturally as a quaternionic spacetime scalar field itself. It correlates via the "spin up/down" projections with two distinct values of the scalar curvature of the Weyl-Cartan geometry underlying the classical gauge theory. Dirac Theory becomes correlated with a theory of purely geometric features in this model. Since "spin space" transformations can be used on the quaternionic Dirac wavefunction field to generate an entire family of paired curvatures that share common Dirac eigenvalues, that becomes a degeneracy of the theory.

As it stands, this theory should account for a number of atomic phenomena, to the degree that Dirac Theory can do so. It may account for some subatomic phenomena as well. It also seems to predict a form of electromagnetic field quanta. On the other hand, the Dirac Equation is still needed for practical computation of results.

There is no evidence currently of quantization or spatial localization of charge in the sources of the electromagnetic field, nor of quantization or spatial localization of rest masses of the wavefunction field. However, until a better theory of the antisymmetric part of the metric is developed, it's not clear what this formalism has to say about variations in rest masses, or the effects of possible excursions of the value of $\hat{a}$ away from the constant value of $(6\imath)^2 = -36$, possibly into the complex plane.

Appendix: Quaternionic Gauges and Curvatures

The close relationship between equations (22) and (23) suggests that equation (22) might be a direct result of generalizing the theory of reference [1] from the complex numbers to the quaternions. If equations (22) and (2) remain true in that generalization, that would imply that the scalar curvature B is fully quaternionic instead of simply being complex valued as in section III.

Now in fact, it is possible to generalize reference [1] from complex numbers to quaternions, but equation (22) is not an immediate result. Rather, the non-commutativity in the generalized structure produces a more complicated generalization of equation (8) as well as other complications not seen in the earlier work. The negative implication of this is that the Dirac Equation itself fits more naturally into the complex plane than it fits into the extended quaternions. Primarily to illustrate just this point, this appendix will briefly sketch a direct extension of the model of reference [1] to quaternionic gauge transformations instead of complex ones.

1. Basics

The basic rules will remain the same. There must always exist a gauge (or gauges) in which the symmetric part of the metric, $g_{\mu\nu}$, is real. In that gauge, $g_{\mu\nu}$ will be denoted by $\tilde{g}_{\mu\nu}$. The antisymmetric part of the metric is still required to be self-dual (dual proportional), and of the form

$$a_{\mu\nu} = h_{\mu\nu} - {}^*h_{\mu\nu}|Q_3 \tag{A.1}$$

where the " $*$ " indicates the dual of the quantity following it. In this, $h_{\mu\nu}$ must be real in the same gauge in which $g_{\mu\nu}$ is real, a requirement that will have additional importance later. For all these quantities, a " $\tilde{}$ " indicates evaluation in the gauge in which the quantity is real. Note the use of $|(-Q_3)$ in place of the $\imath$ used in this same form in reference [1]. That will maintain a simplifying element of commutativity in this form, and elsewhere in the quaternionic generalizations of equations from the earlier work, as already seen in section II.

Likewise, there must always exist a gauge (or gauges) in which equation (6) generalizes to the form

$$v_\mu = -qA_\mu|Q_3 \tag{A.2}$$

with A_μ being the standard, real, electromagnetic potential. This gauge will generally *not* be the same gauge in which $g_{\mu\nu}$ is real.

Since $g_{\mu\nu}$ (and $h_{\mu\nu}$) is real in some gauge, it is always possible to write

$$g_{\mu\nu} = \tilde{g}_{\mu\nu}\gamma \tag{A.3}$$

where γ is some quaternionic scalar. This allows the various operations with $g_{\mu\nu}$ and $h_{\mu\nu}$ to be set up through easy correspondence with operations on real or complex forms. For example, $g^{\mu\nu}$ is easily calculated through the standard requirement that

$$g^{\mu\alpha}g_{\alpha\nu} = \delta^{\mu}_{\nu} \tag{A.4}$$

where δ^{μ}_{ν} is the Kronecker delta. In fact,

$$g^{\mu\nu} = \gamma^{-1}\tilde{g}^{\mu\nu} \tag{A.5}$$

Since gauge transformations are quaternionic, they can generally be applied to $g_{\mu\nu}$ from either the left or right, as denoted by

$$\bar{g}_{\mu\nu} = \lambda g_{\mu\nu}\rho \tag{A.6}$$

where λ is the left gauge transformation, and ρ is the right. However, more restricted cases where one of these two multipliers is always taken to be 1 will be most useful here. Since equation (A.6) can always be written as

$$\bar{g}_{\mu\nu} = \lambda\tilde{g}_{\mu\nu}\gamma\rho \tag{A.7}$$

according to equation (A.3), and since $\tilde{g}_{\mu\nu}$ commutes with every scalar multiplier, restricting either λ or ρ to be 1 may not be a serious restriction in practice. However whatever scheme is adopted, $h_{\mu\nu}$ will be assumed to gauge transform by the *same* pattern as is adopted for $g_{\mu\nu}$. The same pattern can be shown to hold for ${}^*h_{\mu\nu}$ provided the dual is defined as

$${}^*h_{\mu\nu} = \frac{1}{2}g_{\mu\tau}(-g)^{-1/4}\,\epsilon^{\tau\sigma\alpha\beta}h_{\alpha\beta}(-g)^{-1/4}\,g_{\sigma\nu} \tag{A.8}$$

where in addition to the definition of ${}^*h^{\tau\sigma}$, one can discern in this the symmetric manner always used to raise and lower metric indices. And, since the leap-over operator "|" keeps the Q_3 out of the way to the far right, equation (A.1) then indicates that $a_{\mu\nu}$ will likewise follow the same gauge transformation pattern as $g_{\mu\nu}$ and $h_{\mu\nu}$, provided the gauge functions themselves either contain no leap-over operations, or contain none other than $|Q_3$. This assumption will always be made about the gauge functions here.

2. Left and Right Covariant Derivatives / Christoffel Symbols

Since $g_{\mu\nu}$ can now have a limited quaternionic nature, the expression

$$g_{\mu\nu;\gamma} = g_{\mu\nu,\gamma} - g_{\alpha\nu}\left\{ {\alpha \atop \mu\gamma} \right] - g_{\mu\alpha}\left\{ {\alpha \atop \nu\gamma} \right] = 0 \tag{A.9}$$

is not necessarily the same as

$$g_{\mu\nu;\gamma} = g_{\mu\nu,\gamma} - \left[{\alpha \atop \mu\gamma} \right\} g_{\alpha\nu} - \left[{\alpha \atop \nu\gamma} \right\} g_{\mu\alpha} = 0 \tag{A.10}$$

Thus, equation (A.9) defines the covariant derivative of $g_{\mu\nu}$ with respect to the "right handed Christoffel Symbol" $\left\{ {\alpha \atop \mu\nu} \right]$, while equation (A.10) defines the covariant derivative of $g_{\mu\nu}$ with respect to the "left handed Christoffel Symbol" $\left[{\alpha \atop \mu\nu} \right\}$. Both equations actually define the associated Christoffel Symbols, giving

$$\left\{ {\alpha \atop \mu\nu} \right] = \frac{1}{2}\,g^{\alpha\tau}(g_{\mu\tau,\nu} + g_{\nu\tau,\mu} - g_{\mu\nu,\tau}) \tag{A.11}$$

and

$$\left[{\alpha \atop \mu\nu} \right\} = \frac{1}{2}\,(g_{\mu\tau,\nu} + g_{\nu\tau,\mu} - g_{\mu\nu,\tau})\,g^{\alpha\tau} \tag{A.12}$$

respectively.

Furthermore, equation (A.4) and $\delta^{\mu}_{\nu,\tau} = 0$ give

$$g^{\beta\mu}{}_{,\tau} = -g^{\beta\nu}g_{\nu\alpha,\tau}\,g^{\alpha\mu} \tag{A.13}$$

This and equations (A.9) and (A.10) then give

$$g^{\beta\xi}{}_{;\tau} = g^{\beta\xi}{}_{,\tau} + \left\{ {\beta \atop \alpha\tau} \right] g^{\alpha\xi} + \left\{ {\xi \atop \alpha\tau} \right] g^{\beta\alpha} = 0 \tag{A.14}$$

and

$$g^{\beta\xi}{}_{;\tau} = g^{\beta\xi}{}_{,\tau} + g^{\alpha\xi} \left[{\beta \atop \alpha\tau} \right\} + g^{\beta\alpha} \left[{\xi \atop \alpha\tau} \right\} = 0 \tag{A.15}$$

This indicates that the contravariant metric's indices interact with their associated Christoffel symbols on the opposite side from the covariant metric's indices in the definition of the covariant derivative of the metric. That same convention is now adopted as well here for the form of all covariant derivatives, and also more general affine derivatives, of *any* tensor quantity, where the affine derivative simply uses the more general affine connections ${}_R\Gamma^{\alpha}_{\mu\nu}$ and ${}_L\Gamma^{\alpha}_{\mu\nu}$ in place of the Christoffel Symbols $\left\{ {\alpha \atop \mu\nu} \right]$ and $\left[{\alpha \atop \mu\nu} \right\}$. The more general affine connections are prefixed with the lowered "R" and "L" to maintain different right and left handed forms on the more general level, just as there are right and left handed Christoffel Symbols.

For completeness, now define the " $\tilde{;}$ " derivative as the reversal of the convention just given for the " ; " derivative. That means that all the tensor - Christoffel Symbol positions are reversed for each term for each tensor index in the covariant derivative expression. For example,

$$g_{\mu\nu\tilde{;}\gamma} = g_{\mu\nu,\gamma} - \left\{ {\alpha \atop \mu\gamma} \right] g_{\alpha\nu} - \left\{ {\alpha \atop \nu\gamma} \right] g_{\mu\alpha} \ (\neq 0) \tag{A.16}$$

By definition, the Christoffel Symbols are always defined using a "normal" covariant derivative of the (symmetric part of the) metric tensor. Clearly, tensor - Christoffel Symbol positions in these expressions are determined *both* by the type of Christoffel Symbol ("{]" or "[}"), *and* the use of " ; " or " $\tilde{;}$ " in the derivative.

With these facts established, the "left handed" Christoffel Symbols and more general affine connection ${}_L\Gamma^{\alpha}_{\mu\nu}$ will now be arbitrarily dropped, and the right handed cases used in what follows. However, it should also be noted that the generalization of equation (9) in this extended structure will eventually involve a quaternion conjugate ("QC") term to replace the "CC" term in equation (9). That quaternion conjugate term will tend to involve left handed forms to balance the right handed forms that are now being chosen in the first part of the action. Because of that, the basic left hand should not be suppressed or shortchanged overall, although a somewhat different type of left handedness will also arise at the next level to be examined. That will be found necessarily to enter asymmetrically, but it should still tend to be balanced by corresponding right handed forms in the QC terms in the action.

For general quaternionic M_μ and N_ν,

$$(M_\mu N_\nu)_{;\tau} \neq M_{\mu;\tau} N_\nu + M_\mu N_{\nu;\tau} \tag{A.17}$$

Additionally, contraction on tensor indices inside an already evaluated covariant derivative will not necessarily equal the covariant derivative of the contracted quantity. Examples such as these limit the usefulness of these gauge varying, generalized covariant derivatives outside of gauges in which quantities commute easily in products. However, in other cases, these covariant derivatives will still be helpful, and will be used. On the other hand, any genuine physics in this structure will require gauge invariant constructions, including gauge invariant covariant derivatives. Those will be developed later, and they will involve real Christoffel Symbols that thus avoid these limitations just noted.

3. Weyl-Like Connections, Gauge Properties, and Curvatures

Since right handed forms are now chosen, such as equation (A.11), specialize equation (A.6) to $\lambda = 1$, or

$$\bar{g}_{\mu\nu} = g_{\mu\nu}\rho \tag{A.18}$$

Corresponding to that,

$$\bar{g}^{\mu\nu} = \rho^{-1} g^{\mu\nu} \tag{A.19}$$

These together with equation (A.11) then give that

$$\begin{aligned}\{\bar{{}^{\alpha}_{\mu\nu}}] &= \rho^{-1}\{{}^{\alpha}_{\mu\nu}]\rho + \tfrac{1}{2}\,\delta^{\alpha}_{\mu}\rho^{-1}\rho_{,\nu} + \tfrac{1}{2}\,\delta^{\alpha}_{\nu}\rho^{-1}\rho_{,\mu} - \tfrac{1}{2}\,\rho^{-1}g^{\alpha\tau}g_{\mu\nu}\,\rho_{,\tau} \\ &= \rho^{-1}\{{}^{\alpha}_{\mu\nu}]\rho + \tfrac{1}{2}\,\delta^{\alpha}_{\mu}\rho^{-1}\rho_{,\nu} + \tfrac{1}{2}\,\delta^{\alpha}_{\nu}\rho^{-1}\rho_{,\mu} - \tfrac{1}{2}\,g^{\alpha\tau}g_{\mu\nu}\,\rho^{-1}\rho_{,\tau}\end{aligned} \tag{A.20}$$

The second line follows from the fact that

$$\begin{aligned}g^{\alpha\tau}g_{\mu\nu} &= \gamma^{-1}\tilde{g}^{\alpha\tau}\tilde{g}_{\mu\nu}\gamma \\ &= \tilde{g}^{\alpha\tau}\tilde{g}_{\mu\nu} \\ &= \tilde{g}_{\mu\nu}\tilde{g}^{\alpha\tau} \\ &= \tilde{g}_{\mu\nu}\gamma\gamma^{-1}\tilde{g}^{\alpha\tau} \\ &= g_{\mu\nu}g^{\alpha\tau}\end{aligned} \tag{A.21}$$

using equations (A.3) and (A.5). That metric combination is always real, gauge invariant, and commutes *as a unit* with everything.

In analogy with reference [1], now require a Cartan affine connection ${}_R\Gamma^{\alpha}_{\mu\nu}$ (hereafter denoted by $\Gamma^{\alpha}_{\mu\nu}$) to exist which gauge transforms in analogy to Einstein's "lambda invariance"[26] via

$$\bar{\Gamma}^{\alpha}_{\mu\nu} = \rho^{-1}\Gamma^{\alpha}_{\mu\nu}\,\rho + \tfrac{1}{2}\,\delta^{\alpha}_{\mu}\,\rho^{-1}\rho_{,\nu} \tag{A.22}$$

This is done by defining the connection as

$$\Gamma^{\alpha}_{\mu\nu} = \{{}^{\alpha}_{\mu\nu}] + \delta^{\alpha}_{\nu}v_{\mu} - g^{\alpha\tau}g_{\mu\nu}v_{\tau} \tag{A.23}$$

where v_μ is the quantity of equation (A.2), and

$$\begin{aligned}\bar{v}_{\mu} &= \rho^{-1}(v_{\mu} - \tfrac{1}{2}\,\rho_{,\mu}\rho^{-1})\rho \\ &= \rho^{-1}v_{\mu}\,\rho - \tfrac{1}{2}\,\rho^{-1}\rho_{,\mu}\end{aligned} \tag{A.24}$$

Both of the connections $\Gamma^{\alpha}_{\mu\nu}$ and $\{{}^{\alpha}_{\mu\nu}]$ have associated curvature tensors, but those have both a "right handed" and a "left handed" form themselves, irrespective of the right-left nature of the underlying connection used in them. Specifically,

$${}_RB^{\gamma}_{\mu\tau\sigma} = \Gamma^{\gamma}_{\mu\sigma,\tau} - \Gamma^{\gamma}_{\mu\tau,\sigma} + \Gamma^{\gamma}_{\eta\tau}\Gamma^{\eta}_{\mu\sigma} - \Gamma^{\gamma}_{\eta\sigma}\Gamma^{\eta}_{\mu\tau} \tag{A.25}$$

$${}_LB^{\gamma}_{\mu\tau\sigma} = \Gamma^{\gamma}_{\mu\sigma,\tau} - \Gamma^{\gamma}_{\mu\tau,\sigma} + \Gamma^{\eta}_{\mu\sigma}\Gamma^{\gamma}_{\eta\tau} - \Gamma^{\eta}_{\mu\tau}\Gamma^{\gamma}_{\eta\sigma} \tag{A.26}$$

$${}_RR^{\gamma}_{\mu\tau\sigma} = \{{}^{\gamma}_{\mu\sigma}]_{,\tau} - \{{}^{\gamma}_{\mu\tau}]_{,\sigma} + \{{}^{\gamma}_{\eta\tau}]\{{}^{\eta}_{\mu\sigma}] - \{{}^{\gamma}_{\eta\sigma}]\{{}^{\eta}_{\mu\tau}] \tag{A.27}$$

and

$${}_LR^{\gamma}_{\mu\tau\sigma} = \{{}^{\gamma}_{\mu\sigma}]_{,\tau} - \{{}^{\gamma}_{\mu\tau}]_{,\sigma} + \{{}^{\eta}_{\mu\sigma}]\{{}^{\gamma}_{\eta\tau}] - \{{}^{\eta}_{\mu\tau}]\{{}^{\gamma}_{\eta\sigma}] \tag{A.28}$$

When equation (A.22) is substituted into equations (A.25) and (A.26) to obtain the gauge properties of those curvature tensors, neither curvature form transforms in a particularly neat manner by itself. Surprisingly however, the combination

$$B^{\gamma}_{\mu\tau\sigma} = \tfrac{3}{2}\,{}_RB^{\gamma}_{\mu\tau\sigma} - \tfrac{1}{2}\,{}_LB^{\gamma}_{\mu\tau\sigma} \tag{A.29}$$

does have neat gauge transformation properties (reverse roles of ${}_RB^{\gamma}_{\mu\tau\sigma}$ and ${}_LB^{\gamma}_{\mu\tau\sigma}$ if left handed forms with $\rho = 1$ and $\lambda \neq 1$ are adopted initially rather than right). Specifically

$$\bar{B}^{\gamma}_{\mu\tau\sigma} = \rho^{-1}B^{\gamma}_{\mu\tau\sigma}\,\rho \tag{A.30}$$

which is the same form as the gauge transformation of a Yang-Mills Field in SU(2) gauge theory[18]. Furthermore $B^{\gamma}_{\mu\tau\sigma}$ actually reduces to the curvature tensor of reference [1] as it is restricted to the complex plane in which products like $\Gamma^{\gamma}_{\eta\tau}\Gamma^{\eta}_{\mu\sigma}$ commute internally. In other words, it is the appropriate generalization of the curvature tensor of the earlier reference.

Note that the imbalance between ${}_RB^\gamma_{\mu\tau\sigma}$ and ${}_LB^\gamma_{\mu\tau\sigma}$ in equation (A.29) is fixed by the gauge properties of $\Gamma^\alpha_{\mu\nu}$ given by equation (A.22), so it's fundamental to this structure. It is not adjustable for this Cartan type connection without changing the coefficient of the second term in equation (A.22). Furthermore, if one attempts to use a full Weyl connection analog in place of the Cartan form of equation (A.23) by adding the term $\delta^\alpha_\mu v_\nu$ to the right side of that equation, no suitable generalized curvature $B^\gamma_{\mu\tau\sigma}$ emerges at all. Rather, ${}_RB^\gamma_{\mu\tau\sigma}$ and ${}_LB^\gamma_{\mu\tau\sigma}$ will then attempt to enter more symmetrically in the analog of equation (A.29), and that will lead to the disappearance of the derivatives of $\Gamma^\alpha_{\mu\nu}$ from the result. Without the derivative terms, this quantity cannot reduce to the theory of reference [1] in the complex plane at all. In other words, this structure actually discriminates against the analog of Weyl's original theory[27] and favors the Cartan case, which has torsion rather than nonmetricity. This differs from reference [1], in which both the Weyl and Cartan type connections were in principle allowed.

Finally, the four vector v_μ of equation (A.24) has its own directly associated Yang-Mills field tensor[18]. The gauge properties give that

$$\begin{aligned} y_{\mu\nu} &= v_{\nu,\mu} - v_{\mu,\nu} + 2(v_\nu v_\mu - v_\mu v_\nu) \\ &= p_{\mu\nu} + 2(v_\nu v_\mu - v_\mu v_\nu) \end{aligned} \tag{A.31}$$

gauge transforms just as $B^\gamma_{\mu\tau\sigma}$ does, or

$$\bar{y}_{\mu\nu} = \rho^{-1} y_{\mu\nu}\,\rho \tag{A.32}$$

Since by assumption, there exists a gauge in which v_μ has the form of equation (A.2), then in that same gauge

$$y_{\mu\nu} = -qF_{\mu\nu}|Q_3 \tag{A.33}$$

where $F_{\mu\nu}$ is the standard Maxwell tensor, the curl of A_μ, and it is real. But then by equation (A.32), this tensor has this form in *all* gauges provided the gauge function ρ satisfies the assumptions made after equation (A.8), that it either contains no leap-over operations, or contains none other than $|Q_3$. Other than that, ρ is fully quaternionic, but the $|Q_3$ stays out of the way of ρ^{-1} and ρ, allowing them to cancel.

Thus $y_{\mu\nu}$ is actually gauge invariant, and mathematically it remains in the complex plane as a purely imaginary tensor. The leap-over operator has simplified the model, allowing this sort of gauge transformation of one side of a two sided quaternionic product, while the still unaffected, constant unit imaginary quantity operates on the other side of the same product.

4. The Makeup of the Curvature Tensor

Equation (A.23) can be written

$$\Gamma^\alpha_{\mu\nu} = \{^{\ \alpha}_{\mu\nu}] + U^\alpha_{\mu\nu} \tag{A.34}$$

where

$$U^\alpha_{\mu\nu} = \delta^\alpha_\nu v_\mu - g^{\alpha\tau} g_{\mu\nu} v_\tau \tag{A.35}$$

Substituting these into equations (A.25) and (A.26) then gives the surprisingly neat results

$${}_RB^\gamma_{\mu\tau\sigma} = {}_RR^\gamma_{\mu\tau\sigma} + U^\gamma_{\mu\sigma;\tau} - U^\gamma_{\mu\tau;\sigma} + U^\gamma_{\eta\tau}U^\eta_{\mu\sigma} - U^\gamma_{\eta\sigma}U^\eta_{\mu\tau} \tag{A.36}$$

and

$${}_LB^\gamma_{\mu\tau\sigma} = {}_LR^\gamma_{\mu\tau\sigma} + U^\gamma_{\mu\sigma\tilde{;}\tau} - U^\gamma_{\mu\tau\tilde{;}\sigma} + U^\eta_{\mu\sigma}U^\gamma_{\eta\tau} - U^\eta_{\mu\tau}U^\gamma_{\eta\sigma} \tag{A.37}$$

These equations are one example (perhaps the best) in which the " ; " and " $\tilde{;}$ " covariant derivatives give results that are both compact, and express useful information.

However, in order to proceed further with the evaluation of equations (A.36) and (A.37), the covariant derivatives should be written out as partial derivatives and Christoffel Symbol terms, and equation (A.35) should be substituted into the result. Furthermore, any resulting partial derivatives of $g_{\mu\nu}$ or $g^{\mu\nu}$ should be evaluated using equations (A.9) and (A.14) to substitute terms with Christoffel Symbols for the partial derivative terms. In practice, the combination $(g^{\gamma\eta}g_{\mu\sigma})_{,\tau}$ always appears as a unit, and can be eliminated using

$$\begin{aligned} (g^{\gamma\eta}g_{\mu\sigma})_{,\tau} &= g^{\gamma\eta}g_{\alpha\sigma}\{^{\ \alpha}_{\mu\tau}] + g^{\gamma\eta}g_{\mu\alpha}\{^{\ \alpha}_{\sigma\tau}] \\ &\quad -\{^{\ \gamma}_{\alpha\tau}]g^{\alpha\eta}g_{\mu\sigma} - \{^{\ \eta}_{\alpha\tau}]g^{\gamma\alpha}g_{\mu\sigma} \end{aligned} \tag{A.38}$$

keeping equation (A.21) in mind for the result. The fact that the $g^{\gamma\eta}g_{\mu\sigma}$ terms are real then allows equation (A.38) to have more than one valid form simply by varying the position of such terms in its products. However, the same

form should be chosen throughout either ${}_{R}B^{\gamma}_{\mu\tau\sigma}$ or ${}_{L}B^{\gamma}_{\mu\tau\sigma}$ internally to avoid possibly encountering extraneous terms that should evaluate to zero with some effort, but are more easily avoided from the outset through consistency. Additionally, the full expression of equation (A.38) itself should be real, and could be moved around as a unit in products in its containing equation if necessary. However, all this flexibility leads to more than one expansion of $B^{\gamma}_{\mu\tau\sigma}$ in gauge varying quantities like v_{μ}, although all the expansions are equivalent, and all will lead to the same, unique, gauge invariant result in what follows. Since the gauge invariant result contains any real physics, its uniqueness is what is important.

The result of the above substitutions and expansions gives

$$
\begin{aligned}
B^{\gamma}_{\mu\tau\sigma} \;=\; & \tfrac{3}{2}\,{}_{R}R^{\gamma}_{\mu\tau\sigma} - \tfrac{1}{2}\,{}_{L}R^{\gamma}_{\mu\tau\sigma} \\
& +(\tfrac{3}{2}\,v_{\mu;\tau} - \tfrac{1}{2}\,v_{\mu\tilde{;}\tau})\delta^{\gamma}_{\sigma} - (\tfrac{3}{2}\,v_{\mu;\sigma} - \tfrac{1}{2}\,v_{\mu\tilde{;}\sigma})\delta^{\gamma}_{\tau} \\
& -(\tfrac{3}{2}\,v_{\eta;\tau} - \tfrac{1}{2}\,v_{\eta\tilde{;}\tau})g^{\eta\gamma}g_{\mu\sigma} + (\tfrac{3}{2}\,v_{\eta;\sigma} - \tfrac{1}{2}\,v_{\eta\tilde{;}\sigma})g^{\eta\gamma}g_{\mu\tau} \\
& +2\,[v_{\eta}, \left\{{}^{\;\gamma}_{\alpha\tau}\right]]\,g^{\eta\alpha}g_{\mu\sigma} - 2\,[v_{\eta}, \left\{{}^{\;\gamma}_{\alpha\sigma}\right]]\,g^{\eta\alpha}g_{\mu\tau} \\
& +(\tfrac{3}{2}\,v_{\sigma}v_{\mu} - \tfrac{1}{2}\,v_{\mu}v_{\sigma})\delta^{\gamma}_{\tau} - (\tfrac{3}{2}\,v_{\tau}v_{\mu} - \tfrac{1}{2}\,v_{\mu}v_{\tau})\delta^{\gamma}_{\sigma} \\
& -v_{\eta}v_{\beta}g^{\eta\beta}g_{\mu\sigma}\delta^{\gamma}_{\tau} + v_{\eta}v_{\beta}g^{\eta\beta}g_{\mu\tau}\delta^{\gamma}_{\sigma} \\
& +(\tfrac{3}{2}\,v_{\alpha}v_{\tau} - \tfrac{1}{2}\,v_{\tau}v_{\alpha})g^{\alpha\gamma}g_{\mu\sigma} - (\tfrac{3}{2}\,v_{\alpha}v_{\sigma} - \tfrac{1}{2}\,v_{\sigma}v_{\alpha})g^{\alpha\gamma}g_{\mu\tau}
\end{aligned}
\tag{A.39}
$$

where the "[,]" terms are conventional commutators. Those commutators will clearly vanish in gauges in which $\left\{{}^{\;\gamma}_{\alpha\sigma}\right]$ is real.

This can be contracted to give

$$B_{\mu\tau} = B^{\omega}_{\mu\tau\omega} \tag{A.40}$$

and clearly equation (A.30) gives that

$$\bar{B}_{\mu\tau} = \rho^{-1}B_{\mu\tau}\,\rho \tag{A.41}$$

The similarity between this equation and equation (A.32) might then raise expectations that the antisymmetric part of $B_{\mu\tau}$ will be proportional to $y_{\mu\tau}$. However, this is not quite the case. A check reveals that the antisymmetric part of $\frac{3}{2}\,{}_{R}R^{\omega}_{\mu\tau\omega} - \frac{1}{2}\,{}_{L}R^{\omega}_{\mu\tau\omega}$ equals $\frac{1}{2}\,[\gamma^{-1}\gamma_{,\mu}, \gamma^{-1}\gamma_{,\tau}]$ where γ is the gauge function in $g_{\mu\nu} = \tilde{g}_{\mu\nu}\gamma$. Since this commutator does not generally vanish, then the antisymmetric part of $\frac{3}{2}\,{}_{R}R^{\omega}_{\mu\tau\omega} - \frac{1}{2}\,{}_{L}R^{\omega}_{\mu\tau\omega}$ is generally not zero, and that antisymmetric component must be balanced elsewhere by antisymmetric gauge terms, even though it vanishes in gauges in which $g_{\mu\nu}$ is real, and also when γ remains in the complex plane. The consequences of this will become clearer in the next section.

Finally define the scalar curvature

$$B = B_{\mu\tau}M^{\mu\tau} \tag{A.42}$$

where $M^{\mu\tau}$ is the full asymmetric inverse metric. It is given by

$$M^{\mu\tau} = (g^{\mu\tau} - a^{\mu\tau})(1 + \tfrac{1}{4}\,a)^{-1} \tag{A.43}$$

where $a = a_{\mu\nu}a^{\mu\nu}$, a gauge invariant quantity which actually lies in the complex plane still, since it is real except for including a $|Q_3$. Since

$$\bar{M}^{\mu\tau} = \rho^{-1}M^{\mu\tau} \tag{A.44}$$

then equations (A.41) and (A.42) give

$$\bar{B} = \rho^{-1}B \tag{A.45}$$

This is then the key quantity needed to define gauge invariant variables. As conceived by Weyl and Eddington[27, 28], it is basically an intrinsic yardstick provided by the spacetime structure itself to reduce equations to dimensionless, gauge invariant quantities that can correspond to actual physics. For this purpose, it is assumed to be nonzero.

5. Gauge Invariant Variables and Their Fundamental Identity

Now by equations (A.18) and (A.45), equation (4) generalizes immediately to

$$\hat{g}_{\mu\nu} = g_{\mu\nu}(B/C) \tag{A.46}$$

and its inverse

$$\hat{g}^{\mu\nu} = CB^{-1}g^{\mu\nu} \tag{A.47}$$

for the symmetric part of the gauge invariant metric. Indeed, the full asymmetric metric has an analogous form as

$$\begin{aligned} \hat{m}_{\mu\nu} &= m_{\mu\nu}(B/C) \\ &= \hat{g}_{\mu\nu} + \hat{a}_{\mu\nu} \end{aligned} \tag{A.48}$$

These are all gauge invariant, and all are required to be real except for the $|Q_3$ multiplying ${}^*\hat{h}_{\mu\nu}$ in the gauge invariant version of equation (A.1). The full inverse asymmetric metric of equations (A.43) and (A.44) clearly generalizes to

$$\begin{aligned} \hat{M}^{\mu\nu} &= CB^{-1}M^{\mu\nu} \\ &= (\hat{g}^{\mu\tau} - \hat{a}^{\mu\tau})(1 + \tfrac{1}{4}\,\hat{a})^{-1} \end{aligned} \tag{A.49}$$

where as noted after equation (A.43), $\hat{a}$ is a gauge invariant quantity which actually lies in the complex plane still, since it is real except for including a $|Q_3$.

There is a gauge invariant $\{\hat{{}^{\alpha}_{\mu\nu}}\}$ based on the real $\hat{g}_{\mu\nu}$, and it is real,

$$\{\hat{{}^{\alpha}_{\mu\nu}}\} = \frac{1}{2}\,\hat{g}^{\alpha\tau}(\hat{g}_{\mu\tau,\nu} + \hat{g}_{\nu\tau,\mu} - \hat{g}_{\mu\nu,\tau}) \tag{A.50}$$

without a right-left nature any longer. The covariant derivative with respect to it is indicated by "$_{\|}$", and it is now quite well behaved, including obeying the product rule since $\{\hat{{}^{\alpha}_{\mu\nu}}\}$ commutes with everything. Both the "$\tilde{;}$" and the " ; " covariant derivative conventions will reduce to it. If equations (A.46) and (A.47) are substituted into equation (A.50), the result gives

$$\begin{aligned} \{\hat{{}^{\alpha}_{\mu\nu}}\} &= B^{-1}\{{}^{\alpha}_{\mu\nu}]B + \tfrac{1}{2}\,\delta^{\alpha}_{\mu}B^{-1}B_{,\nu} + \tfrac{1}{2}\,\delta^{\alpha}_{\nu}B^{-1}B_{,\mu} - \tfrac{1}{2}\,B^{-1}g^{\alpha\tau}g_{\mu\nu}\,B_{,\tau} \\ &= B^{-1}\{{}^{\alpha}_{\mu\nu}]B + \tfrac{1}{2}\,\delta^{\alpha}_{\mu}B^{-1}B_{,\nu} + \tfrac{1}{2}\,\delta^{\alpha}_{\nu}B^{-1}B_{,\mu} - \tfrac{1}{2}\,g^{\alpha\tau}g_{\mu\nu}\,B^{-1}B_{,\tau} \end{aligned} \tag{A.51}$$

These real, commuting forms now give us a normal, real, gauge invariant Riemannian geometry on which we can impose a standard form of General Relativity. They define $\hat{R}^{\gamma}_{\mu\tau\sigma}$ as the Riemann Curvature Tensor, and both the conventions of ${}_R R^{\gamma}_{\mu\tau\sigma}$ and ${}_L R^{\gamma}_{\mu\tau\sigma}$ reduce to this same tensor. There is a (now symmetric) $\hat{R}_{\mu\tau} = \hat{R}^{\omega}_{\mu\tau\omega}$, and a scalar $\hat{R} = \hat{g}^{\mu\tau}\hat{R}_{\mu\tau}$.

By equations (A.24) and (A.45), equation (5) generalizes to a gauge invariant

$$\begin{aligned} \hat{v}_{\mu} &= B^{-1}(v_{\mu} - \tfrac{1}{2}\,B_{,\mu}B^{-1})B \\ &= B^{-1}v_{\mu}B - \tfrac{1}{2}\,B^{-1}B_{,\mu} \end{aligned} \tag{A.52}$$

which is fully quaternionic generally. Then in analogy to equation (A.23), define the gauge invariant

$$\hat{\Gamma}^{\alpha}_{\mu\nu} = \{\hat{{}^{\alpha}_{\mu\nu}}\} + \delta^{\alpha}_{\nu}\hat{v}_{\mu} - \hat{g}^{\alpha\tau}\hat{g}_{\mu\nu}\hat{v}_{\tau} \tag{A.53}$$

Substituting into this from equations (A.51) and (A.52), and using equation (A.23), one sees

$$\hat{\Gamma}^{\alpha}_{\mu\nu} = B^{-1}\Gamma^{\alpha}_{\mu\nu}\,B + \tfrac{1}{2}\,\delta^{\alpha}_{\mu}\,B^{-1}B_{,\nu} \tag{A.54}$$

But this is exactly the same form as a gauge transformation on $\Gamma^{\alpha}_{\mu\nu}$ as defined in equation (A.22). Thus, if one defines ${}_R\hat{B}^{\gamma}_{\mu\tau\sigma}$ and ${}_L\hat{B}^{\gamma}_{\mu\tau\sigma}$ using $\hat{\Gamma}^{\alpha}_{\mu\nu}$ in full analogy to the use of $\Gamma^{\alpha}_{\mu\nu}$ in ${}_R B^{\gamma}_{\mu\tau\sigma}$ and ${}_L B^{\gamma}_{\mu\tau\sigma}$, the result gives finally that

$$\begin{aligned} \hat{B}^{\gamma}_{\mu\tau\sigma} &= \tfrac{3}{2}\,{}_R\hat{B}^{\gamma}_{\mu\tau\sigma} - \tfrac{1}{2}\,{}_L\hat{B}^{\gamma}_{\mu\tau\sigma} \\ &= B^{-1}B^{\gamma}_{\mu\tau\sigma}B \end{aligned} \tag{A.55}$$

This can be expanded just like equation (A.39), but now with so many quantities real, the much simpler result is

$$\begin{aligned} \hat{B}^{\gamma}_{\mu\tau\sigma} = \;& \hat{R}^{\gamma}_{\mu\tau\sigma} + \hat{v}_{\mu\|\tau}\delta^{\gamma}_{\sigma} - \hat{v}_{\mu\|\sigma}\delta^{\gamma}_{\tau} \\ & -\hat{v}_{\eta\|\tau}\hat{g}^{\eta\gamma}\hat{g}_{\mu\sigma} + \hat{v}_{\eta\|\sigma}\hat{g}^{\eta\gamma}\hat{g}_{\mu\tau} \\ & +(\tfrac{3}{2}\,\hat{v}_{\sigma}\hat{v}_{\mu} - \tfrac{1}{2}\,\hat{v}_{\mu}\hat{v}_{\sigma})\delta^{\gamma}_{\tau} - (\tfrac{3}{2}\,\hat{v}_{\tau}\hat{v}_{\mu} - \tfrac{1}{2}\,\hat{v}_{\mu}\hat{v}_{\tau})\delta^{\gamma}_{\sigma} \\ & -\hat{v}_{\eta}\hat{v}_{\beta}\hat{g}^{\eta\beta}\hat{g}_{\mu\sigma}\delta^{\gamma}_{\tau} + \hat{v}_{\eta}\hat{v}_{\beta}\hat{g}^{\eta\beta}\hat{g}_{\mu\tau}\delta^{\gamma}_{\sigma} \\ & +(\tfrac{3}{2}\,\hat{v}_{\alpha}\hat{v}_{\tau} - \tfrac{1}{2}\,\hat{v}_{\tau}\hat{v}_{\alpha})\hat{g}^{\alpha\gamma}\hat{g}_{\mu\sigma} - (\tfrac{3}{2}\,\hat{v}_{\alpha}\hat{v}_{\sigma} - \tfrac{1}{2}\,\hat{v}_{\sigma}\hat{v}_{\alpha})\hat{g}^{\alpha\gamma}\hat{g}_{\mu\tau} \end{aligned} \tag{A.56}$$

Much of the right-left distinction of equation (A.39), along with the commutators, is now gone. The main left-right distinction remaining is in the terms involving products of $\hat{v}_{\mu}$, because that quantity is fully quaternionic still.

This equation and equation (A.55) then contract to give

$$\begin{aligned}
\hat{B}_{\mu\tau} &= \hat{B}^{\omega}_{\mu\tau\omega} \\
&= B^{-1} B_{\mu\tau} B \\
&= \hat{R}_{\mu\tau} + (\hat{v}_{\mu\|\tau} + \hat{v}_{\tau\|\mu}) + \hat{v}^{\alpha}{}_{\|\alpha} \hat{g}_{\mu\tau} \\
&\quad -(\hat{v}_{\mu}\hat{v}_{\tau} + \hat{v}_{\tau}\hat{v}_{\mu}) + 2\hat{v}^{\alpha}\hat{v}_{\alpha}\hat{g}_{\mu\tau} \\
&\quad +(\hat{v}_{\mu\|\tau} - \hat{v}_{\tau\|\mu}) + 4(\hat{v}_{\mu}\hat{v}_{\tau} - \hat{v}_{\tau}\hat{v}_{\mu})
\end{aligned} \tag{A.57}$$

where the symmetric and antisymmetric parts have been clearly separated with the antisymmetric part all on the last line. Because $\hat{v}_{\mu\|\tau} - \hat{v}_{\tau\|\mu} = \hat{v}_{\mu,\tau} - \hat{v}_{\tau,\mu}$, it is

$$\begin{aligned}
-\hat{w}_{\mu\tau} &= \hat{v}_{\mu\|\tau} - \hat{v}_{\tau\|\mu} + 4(\hat{v}_{\mu}\hat{v}_{\tau} - \hat{v}_{\tau}\hat{v}_{\mu}) \\
&= \hat{v}_{\mu,\tau} - \hat{v}_{\tau,\mu} + 4(\hat{v}_{\mu}\hat{v}_{\tau} - \hat{v}_{\tau}\hat{v}_{\mu}) \\
&= -\hat{y}_{\mu\tau} - 2(\hat{v}_{\tau}\hat{v}_{\mu} - \hat{v}_{\mu}\hat{v}_{\tau})
\end{aligned} \tag{A.58}$$

where

$$\begin{aligned}
\hat{y}_{\mu\tau} &= \hat{v}_{\tau,\mu} - \hat{v}_{\mu,\tau} + 2(\hat{v}_{\tau}\hat{v}_{\mu} - \hat{v}_{\mu}\hat{v}_{\tau}) \\
&= B^{-1} y_{\mu\tau} B
\end{aligned} \tag{A.59}$$

by equations (A.31) and (A.32), since equation (A.52) has the same form as the gauge transformation of v_μ in equation (A.24).

Provided the quaternionic value of B contains no leap-over operations, or contains none other than $|Q_3$, equation (A.33) implies that

$$\hat{y}_{\mu\nu} = -qF_{\mu\nu}|Q_3 \tag{A.60}$$

in which the gauge $v_\mu = -qA_\mu|Q_3$ is now implicit. But then,

$$\begin{aligned}
\hat{w}_{\mu\nu} &= -qF_{\mu\nu}|Q_3 + \frac{1}{2}\left[B^{-1}B_{,\nu}, B^{-1}B_{,\mu}\right] \\
&= \hat{p}_{\mu\nu} - \frac{1}{2}\left[B^{-1}B_{,\mu}, B^{-1}B_{,\nu}\right]
\end{aligned} \tag{A.61}$$

Clearly $\hat{w}_{\mu\nu}$ has a curvature generated, quaternionic tail on it, unlike $\hat{y}_{\mu\nu}$, which remains in the complex plane. This was anticipated above when the antisymmetric part of $\frac{3}{2}\,{}_{R}R^{\omega}_{\mu\tau\omega} - \frac{1}{2}\,{}_{L}R^{\omega}_{\mu\tau\omega}$ was noted to require additional antisymmetric gauge terms to gauge balance it. This is the form those extra terms take in $\hat{w}_{\mu\nu}$.

Given that $B = B_{\mu\tau}M^{\mu\tau}$, $\hat{M}^{\mu\nu} = CB^{-1}M^{\mu\nu}$, and $\hat{B}_{\mu\tau} = B^{-1}B_{\mu\tau}B$, then

$$\begin{aligned}
\hat{B} &= \hat{B}_{\mu\tau}\hat{M}^{\mu\nu} \\
&= CB^{-1}B_{\mu\tau}M^{\mu\nu} \\
&= C
\end{aligned} \tag{A.62}$$

This is the fundamental, kinematic identity the gauge invariant variables must satisfy by virtue of their definitions, and the geometry's kinematics. Substituting from equations (A.49) and (A.57) for $\hat{M}^{\mu\tau}$ and the expansion of $\hat{B}_{\mu\tau}$, gives this as

$$\hat{R} + 6\hat{v}^{\mu}{}_{\|\mu} + 6\hat{v}^{\mu}\hat{v}_{\mu} + \hat{a}^{\mu\nu}\hat{w}_{\mu\nu} = C(1 + \tfrac{1}{4}\,\hat{a}) \tag{A.63}$$

This is almost exactly the same form as the identity in the complex plane, given in equation (8). However now, the quaternionic $\hat{w}_{\mu\nu}$ replaces $\hat{p}_{\mu\nu}$, and $\hat{v}_\mu$ is fully quaternionic. In both cases, it is only the quaternionic nature of B that pushes these out of the complex plane, and into the full quaternions. Nevertheless, these are significant complications when compared to the version in the complex plane.

Most significantly, the basic Ricatti Equation change of variable, $B = \psi^{-2}$ used in all the earlier versions of this model[1, 29], no longer linearizes equation (A.63) into the "wave equation" without restricting B back into the complex plane. A yet unknown generalization of this change of variable is needed at the very least, one that reduces to $B = \psi^{-2}$ as variables are restricted to the complex plane. This alone is a major complication. Without the basic

"wave equation" form, there is no Dirac Equation, even if the square root in the spin term extends into the quaternions as the Dirac spin term must.

Beyond that, a straightforward analogy to the value of $\hat{a}^{\mu\nu}$ obtained in reference [1] would suggest that here,

$$\hat{a}^{\mu\nu} = K\left(\sqrt{\hat{W}}\right)^{-1}\hat{W}^{\mu\nu} \tag{A.64}$$

where

$$\hat{W}_{\mu\nu} = \hat{w}_{\mu\nu} - {}^{*}\hat{w}_{\mu\nu}|Q_3 \tag{A.65}$$

and

$$\hat{W} = \hat{W}^{\mu\nu}\hat{W}_{\mu\nu} \tag{A.66}$$

But those results all assume $\hat{w}_{\mu\nu}$ itself is in the complex plane. When it isn't, such forms would require $\hat{h}_{\mu\nu}$ in $\hat{a}_{\mu\nu}$ not to be real, in contradiction to the original assumptions about the form of $\hat{a}_{\mu\nu}$ in equation (A.1). Indeed, the more general quaternionic nature of $\hat{W}_{\mu\nu}$ wouldn't even allow an $\hat{h}_{\mu\nu}$ that's a real antisymmetric tensor to within some scalar quaternionic multiplying factor. And the alternative of attempting to generalize the structure of the antisymmetric part of the metric $\hat{a}_{\mu\nu}$ to accommodate elements with so much independent quaternion nature, element to element, also creates major complications. It would certainly be a more general geometric structure than originally envisioned. Accordingly, I see no alternative but to add a new constraint to directly constrain $\hat{h}_{\mu\nu}$ to be real in the action (which is otherwise constructed in analogy to equation (9)). However, even if this is done, the spin term in equation (A.63) should still not entirely lose its more general quaternionic tail generated in equation (A.61). That tail explains why the curvature solutions for B_u and B_d in equations (50) and (51) are considered separate complex values, not reassembled into a quaternion. A quaternion would generally spawn that tail, differing from Dirac Theory.

Clearly for both these reasons then, equation (A.63) does *not* give the Dirac Equation, but rather some more general condition that's more complicated in form. Thus, even though equation (26) suggests that the Dirac spin term is a natural result in quaternions, equation (22) still has more in common with the complex plane. At least in this formalism, a deeper excursion into quaternions forces the electromagnetic field to generalize into a fully quaternionic "hyper Yang-Mills field", and it makes it more difficult to obtain a linear equation for a wavefunction. Thus if the Dirac Equation is to be found in this approach, it has to be found largely in the complex plane still, as it is in fact seen to be in section III. If this extended structure has interest for any physics, then that would appear necessarily to involve phenomena beyond Dirac Theory. That discussion is reserved for a future work.

Acknowledgments

I would like to thank two members of the audience at my presentation of the Rankin-Taylor contributed talk at the April, 1997 APS Spring Meeting[30], as well as Engelbert Schücking, all for suggesting that I look at Clifford Algebra generalizations of reference [1]. Additionally, I would like to thank Jim Wheeler, David Finkelstein, Daniel Galehouse, Egon Marx, and James Anderson for helpful discussions and points that have contributed significantly to this paper. I would also like to thank Frank Taylor, who coauthored several conference contributions on this theory with me. And, I would like to thank Wayne Bowers, Thomas J. Boyd, Jr., and Bob Levine for helpful discussions and general support during various stages of this work.

---

[1] J. E. Rankin, *Class. Quantum Grav.* **9**, 1045 (1992).
[2] J. E. Rankin, *Preprint RCTP9401*, arXiv:gr-qc/9404023v3 (1994).
[3] J. E. Rankin, published in *Proceedings of the Seventh Grossmann Meeting.* Also available as arXiv:gr-qc/9408029v2.
[4] My thanks to Daniel Galehouse for correctly insisting that cases which correspond to negative values of B (which thus require $C = -1$) are legitimate.
[5] D. C. Galehouse, *Int. J. Theor. Phys.* **20**, 457 (1981).
[6] D. C. Galehouse, *Int. J. Theor. Phys.* **20**, 787 (1981).
[7] D. Finkelstein, J.M. Jauch, S. Schiminovich and D. Speiser, *J. Math. Phys.* **3**, 207 (1962).
[8] S. L. Adler, *Quaternionic Quantum Mechanics and Quantum Fields* (Oxford, 1995).
[9] L.I. Schiff, *Quantum Mechanics* (McGraw-Hill, 1968).
[10] P.M. Morse, H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, 1953).

[11] S. De Leo and P. Rotelli, *Mod.Phys.Lett.* **A11**, 357 (1996). Also available as arXiv:hep-th/9509059v1.
[12] C. Schwartz, *J. Math. Phys.* **48**, 052303 (2007).
[13] J. J. Sakurai, *Advanced Quantum Mechanics* (Reading, MA: Addison-Wesley, 1967).
[14] W. L. Bade and H. Jehle, *Rev. Mod. Phys.* **25**, 714 (1953).
[15] My thanks to Jim Wheeler for making these observations to me privately.
[16] K. R. Symon, *Mechanics* (Reading, MA: Addison-Wesley, 1960).
[17] A. Doroshkevich, J. Hansen, D. Novikov, I. Novikov, D. H. Park, A. Shatskiy, *Phys. Rev. D* **81**, 124011 (2010). Also available as arXiv:0908.1300v3 [gr-qc].
[18] M. Carmeli, *Classical Fields, General Relativity and Gauge Theory* (New York: Wiley, 1982).
[19] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, 1964).
[20] R.M. Wald, *General Relativity* (University of Chicago, 1984).
[21] R. Adler, M. Bazin, M. Schiffer, *Introduction to General Relativity* (McGraw-Hill, 1965).
[22] Separate articles by D.J. Gross and Y. Totsuka, *Texas/PASCOS '92: Relativistic Astrophysics And Particle Cosmology*, C.W. Akerlof and M.A. Srednicki, eds. (Ann. N.Y. Acad. Sci. Vol. 688, 1993). See pp. 160 (Gross), and 348 (Totsuka).
[23] L. Infeld and T. E. Hull, *Rev. Mod. Phys.* **23**, 21 (1951).
[24] W. E. Lamb, Jr. and M. O. Scully, The photoelectric effect without photons, in *Polarization, Matière et Rayonnement* (Presses University de France, Paris, 1969). Also available at `http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19680009569_1968009569.pdf`
[25] J. J. Thorn, M. S. Neel, V. W. Donato, G. S. Bergreen, R. E. Davies, and M. Beck, *Am. J. Phys.* **72** (9), 1210 (2004).
[26] A. Einstein, *The Meaning of Relativity*, (Princeton, 1955).
[27] H. Weyl, *Space-Time-Matter* (Dover, 1922).
[28] A.S. Eddington, *The Mathematical Theory Of Relativity* (Cambridge, 1965).
[29] J. E. Rankin, *Int. J. Theor. Phys.* **20**, 231 (1981).
[30] J. Rankin, F. Taylor, *Bulletin of the American Physical Society*, Session M11, (April, 1997).